INTRODUCTION TO EXCEL

AfterWork

LEARNING OUTCOMES

- I can perform **data cleaning** techniques to prepare data for further analysis.
- I can perform **data analysis** techniques to answer business problems.
- I can perform **data visualisation** techniques to communicate insights to stakeholders.

AfterWork

INTRODUCTION

Spreadsheets
- Most familiar use is data presentation.
- Formulas are useful for computing sales, budgets, and other numeric summaries.
- Convenient for analysis of structured data and data visualisation.

Interesting facts
- A large fraction of the world's structured data is managed and manipulated in spreadsheets.
- Microsoft Excel is dominant tool; many features, but proprietary and expensive.
- Google Sheets is open and free; fewer features.

AfterWork

PRACTICE: DATA CLEANING

Concepts
- Data Cleaning Techniques

Exercise
- Perform the data cleaning techniques outlined in the given guidelines document and save the file for further analysis.
- Data Cleaning guidelines: https://bit.ly/DataCleaningGuidelines

Dataset
- https://bit.ly/DirtyTitanicDataset

AfterWork

PRACTICE: IMPORTING & EXPORTING

Concepts
- Importing data
- Exporting data

Exercise
- Add new column to the left of column F called Fahrenheit.
- Use formula to compute values from celcius column E.
    - Note: Fahrenheit = 1.8C + 32
- Export the data as a CSV file.

Dataset
- https://bit.ly/CitiesDatasetI

AfterWork

ICE BREAKER

Come up with as many questions as possible that can be answered from the clean titanic dataset.

AfterWork

PRACTICE: DATA ANALYSIS

Concepts
- Sorting, Filtering, Aggregation, Joining, Pivot Tables: Restructuring / aggregation / analysis

Exercise
- How many cities in Italy?
- What is the average latitude:
  - Overall?
  - For cities with temperature < 10 ?
  - For cities with temperature > 10 ?
  - For cities where both the city name and the country name end in the letter "a" ?
  - Which are warmer on average – cities in the EU or cities not in the EU?

Dataset
- https://bit.ly/CitiesDatasetI and https://bit.ly/CountriesDatasetI

AfterWork

PRACTICE: DATA ANALYSIS (TITANIC)

Concepts
- Sorting, Filtering, Aggregation, Joining, Pivot Tables: Restructuring / Aggregation / Analysis.

Exercise
- How many passengers sailed for free?
- How many married women over age 50 embarked in Cherbourg?
- How many passenger ages are missing?
- What is the average fare paid by these passengers?
- What is the most common last name among passengers, and how
- many passengers have that last name? What is the average number
- of passengers per last name?
- What is the average fare paid by passengers in the three classes?
- What is the average age of passengers in the three classes?

Dataset
- Clean Titanic Dataset

AfterWork

PRACTICE: DATA VISUALISATION

- Effective Data visualisation
  - Hans Rosling's 200 Countries, 200 Years, 4 Minutes:
    - https://www.youtube.com/watch?v=jbkSRLYSojo
- Suggested charts are often good ones.
- For help while working with charts i.e. Web search or fiddling
- Don't underestimate the power of basic visualizations
- Bar charts, Pie charts, Scatterplots and Sparklines

AfterWork

PRACTICE: BAR CHARTS

- **Concept**
  - **Bar charts:** When one axis is categories and the other is numeric.

- **Exercise**
  - Ten countries with the highest population, bar chart showing populations.

- **Dataset**
  - https://bit.ly/CountriesDatasetI

AfterWork

PRACTICE: PIE CHARTS

- **Concept**
  - **Pie charts:** To compare relative sizes of categories.

- **Exercise**
  - Pie chart showing relative number of cities with negative longitude and positive longitude.

- **Dataset**
  - https://bit.ly/CitiesDatasetI

AfterWork

PRACTICE: SCATTER PLOTS

- **Concept**
  - **Scatter plots:** When both axes are numeric.

- **Exercise**
  - Use a scatter plot to explore whether there is any relationship between the latitude of cities in a country (x-axis) and the population of that country (y-axis).

- **Dataset**
  - https://bit.ly/CitiesDatasetI and https://bit.ly/CountriesDatasetI

AfterWork

PRACTICE: DATA VISUALISATION (TITANIC)

Concepts
- Data visualisation Techniques

Exercise
- How many passengers sailed for free?
- How many married women over age 50 embarked in Cherbourg?
- How many passenger ages are missing?
- What is the average fare paid by these passengers?
- What is the most common last name among passengers, and how
- many passengers have that last name? What is the average number
- of passengers per last name?
- What is the average fare paid by passengers in the three classes?
- What is the average age of passengers in the three classes?

Dataset
- Clean Titanic Dataset

AfterWork